
*Quality Assurance Methodology
Refinement Series*

***Malawi Field Study:
Comparison of Methods
for Assessing Quality of
Health Worker Performance
Related to Management
of Ill Children***

By

Dr. Lynne Miller Franco

Dr. Ciro Franco

Dr. Newton Kumwenda

Dr. Wilfred Nkhoma

August 1996

Preface

The Quality Assurance Project (QAP) was initiated in 1990 to develop and implement sustainable approaches for improving the quality of health care in less developed countries. QAP has two broad objectives: 1) to provide technical assistance in designing and implementing effective strategies for monitoring quality and correcting systemic deficiencies; and 2) to refine existing methods for ensuring optimal quality health care through an applied research program.

The project's Methodology Refinement component is aimed at developing, refining and validating cost-effective measures for improving the quality of health care. This fourth report in the Quality Assurance Methodology Refinement Series measured the level of statistical agreement between four quality assessment methods: observation, exit interviews with patients/caretakers, record review, and interviews with providers. The study examined the strengths, weaknesses and costs of each method in detecting performance problems. Readers interested in the issue of assessment method validity and reliability are encouraged to see also a separate QAP Methodology Refinement report entitled, *Comparative Validity of Three Methods for Assessment of the Quality of Primary Health Care: Guatemala Field Study*. The two studies have important implications for the use of these methods in quality assessments and routine supervision.

The principal investigators in this study would like to thank the following for ensuring the smooth and fruitful implementation of this study:

- the National Programme Managers for ARI, Malaria, and Diarrhoea for their contributions to the design of data collection instruments and training of interviewers/observers;
- the interviewers/observers for their contribution to the design of the instruments and their consistently high quality work during data collection;
- the health care providers at the 14 health facilities in Lilongwe District for their warm welcome and open participation in our study;

-
- the caretakers who willingly provided us with information about their encounters with health workers;
 - the Quality Assurance Project for its support and technical backstopping during this study;
 - the Community Health Sciences Unit/Ministry of Health for their administrative and logistical support; and
 - The United States Agency for International Development for the funding that made this study possible.

The Quality Assurance Project is funded by the U.S. Agency for International Development under Cooperative Agreement DPE-5992-A-00-0050-00 with the Center for Human Services. Collaborating with the Center for Human Services on this project are the Johns Hopkins University School of Hygiene and Public Health and the Academy for Educational Development.

Contents

List of Tables and Figures	6
Executive Summary	7
I. Introduction	11
II. Objectives of the Study	12
III. Methods	12
IV. Results	16
A. Ability of Quality Assessment Methods to Identify Children with Main Symptoms (Cough, Diarrhoea, Fever, and Ear Problems)	16
B. Comparison of Observation, Exit Interviews, Record Review and Provider Interview Data on Management of the Sick Child	17
1. General Assessment Tasks for Assessment of the Sick Child	18
2. Specific Tasks related to Individual Main Symptoms	20
3. Integrated Management of the Sick Child	27
C. Consistency of Provider Behavior	29
D. Cost of Various Methods of Quality Assessment	32
1. Time Required to Use the Methods	32
2. Time Required to See a Single Case of Each Main Symptom	32
3. Time Required to Assess Management of Serious Cases	34
E. Evaluation of Quality Assessment Methods: Strengths and Weaknesses for Obtaining Data on Quality	35
1. Observation of Provider Performance Using a Checklist	35
2. Exit Interviews with Caretakers about Provider Performance	36
3. Reviewing Records to Assess Provider Performance	37
4. Interviews of Providers about their Performance	37
V. Comparison with Data from Other Studies and Conclusions	38
A. Comparison with Results from Other Studies	38
1. Malawi STD Health Facility Survey	38
2. Study of Comparative Validity of Several Methods for Assessing Health Worker Performance in Guatemala	39
B. Conclusions	39
C. Recommendations	40
APPENDIX A: Detailed Comparisons of Observation, Exit Interview, Record Review, and Provider Interview	43
APPENDIX B: Data Collection Instruments	53

Tables and Figures

Table 1: Tasks Assessed for Quality of Performance	14
Table 2: General Assessment of the Child	18
Table 3: Case Management of Cough	21
Table 4: Case Management of Diarrhoea	23
Table 5: Case Management of Malaria	25
Table 6: Integrated Management of the Sick Child	28
Table 7: Time Required to Assess Quality per Case (Ill Child or Provider)	33
Table 8: Number of Ill Children per Hour during the Cold/Dry Season	34
Figure 1: Consistency of Provider Behavior Related to Cough	31
Figure 2: Consistency of Provider Behavior Related to Diarrhoea	31
Figure 3: Consistency of Provider Behavior Related to Fever/Malaria	32

Executive Summary

Introduction

Supervision is a very important process for supporting health workers and helping ensure the quality of services to communities. This study was designed to identify the strengths, weaknesses, and costs of alternative techniques for assessing quality and to develop guidelines for their use in different settings. The study examined four quality assessment methods: observation of provider-patient encounters, exit interviews with patients/caretakers, record review, and interviews with providers.

Data Collection and Analysis

Data were collected at Under Five clinics at 14 health facilities in Lilongwe district in Malawi. The data collection instruments were based on national standards and WHO's algorithm for integrated management of the sick child. 436 provider-patient encounters were observed, composed of 222 cases of cough, 221 cases of fever, 128 cases of diarrhoea, and 27 cases of ear problems. Furthermore, exit interviews were conducted with 426 caretakers, records were reviewed for 362 children, and 30 providers were interviewed. The Kappa statistic, which tests agreement between two measures taken on the same sample, was used to assess agreement between the various quality assessment methods.

Results

Data on tasks related to general assessment of the child and specific management of cough, diarrhoea, and fever all showed generally poor agreement between observation data and record review; the only exception was for treatment, where agreement was excellent. Provider interview data had only fair agreement with observation data on a limited number of items. However, exit interview data had fair to good agreement with observation for many items, particularly those that were concrete activities caretakers could see or hear.

Agreement between observation and exit interview data, when examined for integrated management of children with multiple main symptoms, was fair to good for most tasks related to children with fever and diarrhoea, but was less consistent for children with fever and cough.

Consistency of provider behavior, over the range of patients they were observed managing, varied by main symptom of the child. Provider behavior was least consistent for cough patients, and most consistent for diarrhoea patients, with behavior for fever patients falling in between.

Assessment of the costs in terms of time used to apply the various methods indicated little difference between the methods for assessment of an individual case, with the exception of provider interviews which are more lengthy but require a smaller sample. However, the length of time needed to obtain a series of cases was longer for exit interviews and observation than for record reviews. Furthermore, the ability to assess severe cases through exit interview and observations was limited: only 11% of children with cough were diagnosed as pneumonia, 6% of those with diarrhoea had moderate or severe dehydration, and no cases of severe malaria were seen.

Conclusions

For many case management tasks, exit interviews with mothers can provide reliable data about what happened in the provider-patient encounter, while provider interviews generally do not provide reliable data about what providers do, and record reviews (based on patient registers) furnish very limited information. In addition, providers are not very consistent in their management of individual patients, implying that a single observation would not be sufficient to draw reliable conclusions.

The costs of observation and exit interview are similar in time, while record reviews are shorter and more easily scheduled. Provider interviews appear to be the least expensive.

The tools used during this study could be improved in the following manner:

1. Discussions with providers during observation would enhance data on classification and diagnosis.
2. Examination of children during the exit interview would provide data on whether the provider had made a correct diagnosis, as well as aiding the mother to recognize certain physical examination tasks.
3. Adding case studies to the provider interview would allow assessment of integrated management of an ill child with multiple main symptoms.

Recommendations

From the data presented in this study, it appears that a hierarchy of methods should be applied by supervisors to assess the pre-conditions for their supervisees' routine correct performance: knowledge, skills, application. Quality assessment by supervisors should start with a combination of provider interviews to assess whether they have the knowledge necessary to conduct proper case management and record review to assess whether they use the correct drugs and dosages for the diagnosis they ascribe. If knowledge is found lacking, then supervisors could either provide on-the-job training or arrange for formal training.

Once supervisors are confident in their supervisees' knowledge, it is appropriate to assess their skills, through direct observation of performance. By using these occasions to discuss with supervisees, supervisors can correct deficiencies in performance. Supervisors would need to spend enough time with providers to observe multiple cases being managed.

When supervisors feel confident about their supervisees' ability to correctly manage cases in their presence, occasional assessment using exit interviews with mothers without the knowledge of the supervisee would allow supervisors to know what is routinely being done. Problems in case management arising here would most probably be due to organizational or motivational causes, since knowledge and skill deficiencies would have already been corrected.

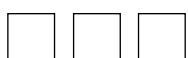
There are several areas requiring exploration or further research:

1. Determining the minimum number of observations/exit interviews needed to make a reliable judgment about provider performance;
2. Improving the provider interview to better accommodate assessment of integrated management of the ill child; and
3. Determining the best assessment tool design; i.e., use of detailed checklists which specify all the tasks for each symptom compared with generic checklists in terms of their ability to identify problems in integrated case management.

Malawi Field Study:

Comparison of Methods for Assessing Quality of Health Worker Performance Related to Management of Ill Children

I. Introduction



Supervision is a very important process for supporting health workers and helping ensure the quality of services to communities. In developing countries, resources for supervision are often limited, and studies have shown that supervisors tend to focus on logistics and personnel issues more than on technical quality.¹ This study was designed to increase knowledge on how best to assess the quality of care, while adding to the meager literature on the cost-effectiveness of various methods for assessing technical performance of health workers. In addition, this study explores some of the issues in supervising integrated management of the sick child, based on the WHO algorithm.

¹ Franco, *Malaria Treatment: A Review of Experiences in Four Countries*, Service Quality Assessment Series, PRICOR (1991: Bethesda, MD); Bums, Franco, and Newman, *Oral Rehydration Therapy in Diarrheal Disease Control: A Review of Experience in Eight Countries*, Service Quality Assessment Series, PRICOR (1 990): Bethesda, MD).

II. Objectives of the Study



The goal of this study was to identify the strengths, weaknesses, and costs of alternative techniques for quality assessment, and, if possible, to develop guidelines for use in different settings. These guidelines should assist supervisors in deciding how best to apply these techniques during routine supervision.

This study examined four assessment methods: observation of providers, exit interviews with caretakers/patients, record review, and interviews with providers. Direct observation was considered for the purpose of this study to be the gold standard; i.e., the measure of “true” performance of the health worker. The major research objective of the study was thus to measure the level of agreement between direct observation and the other three methods. Each method’s effectiveness was also measured in terms of the types of performance problems it can identify and its suitability for assessing infrequently performed tasks (those unlikely to be observable on any specific day). The cost of applying each method was also measured in terms of the time and other resources necessary for routine administration.

III. Methods



This study used cross-sectional, non-experimental methods to compare the results of various assessment methods applied to the same patient-provider encounter. Data from observation, exit interviews, and record review were collected on patient-provider encounters involving at least one of the four major causes of child consultations at health centers and rural hospitals: fever, cough, diarrhoea, and ear problems. In addition, providers were interviewed about their management of such cases.

A total of 436 provider-patient encounters were observed, including 222 cases of cough, 221 cases of fever, 128 cases of diarrhoea, and 27 cases of ear problems. Of the 436 children, 32% had two or more of these main symptoms. Exit interviews were conducted with 426 caretakers, record reviews were conducted on 362 children, and a total of 30 providers were interviewed. The lower number of record reviews reflects the incompleteness of record keeping, with 18% of children not being registered.

The data collection took place in Lilongwe District of Malawi at 12 health centers and 2 rural hospitals. These sites were chosen because the Malawi Essential Drug Programme was, at the time of the study, testing a patient register which could be the basis for record review.* Observers spent 2-5 days at each facility, depending on the number of providers seeing sick children. At most of these facilities, children under five are seen in the Under Five Clinic, which is run by nurses and provides integrated health education, immunizations, growth monitoring, and curative services. However, very sick children coming in the morning hours or those children coming in the afternoon are generally referred to the medical assistants or clinical officers† in the out-patient department (OPD). Of the 30 providers interviewed, half were nurses.

The variables for data collection followed the basic outline of the WHO algorithm for integrated management of the sick child.² For each syndrome (cough, diarrhoea, fever, and ear problems), a series of specific tasks were compiled from the *Malawi Prescriber's Companion* and the Malawi Standard Treatment Guidelines,[‡] and reviewed by the national programme managers for Malaria, Diarrhea Diseases, and Acute Respiratory Infections (ARI). The final list of health worker tasks which were defined as constituting quality care for each main symptom is presented in Table 1.

Three teams composed of two members each carried out the data collection in the following manner: an observer sat with the provider from the moment he/she started seeing patients until the last patient for the day was seen. A second data collector conducted exit interviews with caretakers of those same patients as they left the provider. On the last day at the facility, the person conducting exit interviews interviewed the provider, while the observer extracted the relevant information from the patient register. Time needed to complete an observation, exit interview, record review, or provider interview

² World Health Organization and UNICEF. *Management of Childhood Illness*. World Health Organization, Geneva, 1995.

* At the time of the study, there were no patients registered at government facilities in the rest of Malawi, only a system for counting the number of patients according to age group and diagnosis. The patient is now being introduced at a national level.

† Medical assistants receive 2-3 years training post-high school, while clinical officers receive 3-4 years training. Medical assistants are generally responsible for work in the OPD, while clinical officers also have surgical and in-patient responsibilities.

Table 1				
Health Worker Tasks Assessed for Quality of Performance				
Component of Case Management	Cough	Diarrhoea	Fever	Ear Problem
1. HISTORY	HW asks about: - duration of cough - difficulty breathing	HW asks about: - duration of diarrhoea - presence of blood/mucus in stools - whether child urinating	HW asks about: - ability to drink - other symptoms (cough, throat, ear pain, diarrhoea) - vomiting - any treatment already taken	HW asks about: - ear discharge - duration of ear discharge - ear pain
2. PHYSICAL EXAM	HW: - counts respiration rate - observes for chest indrawing - checks for stiff neck	HW: - pinches skin - looks for sunken eyes - (if infant) checks for sunken fontanel	HW: - takes temperature - checks for enlarged spleen - takes blood for slide - checks for anemia	HW: - examines inside ear - examines behind ear
3. PHYSICAL EXAM	HW: - classifies child by degree of severity - gives appropriate drug - gives correct dosage - treats fever if present - refers case if severe	HW: - classifies child by degree of dehydration - treats dehydration appropriately - refers case if severe or blood - keeps child under observation (4 hrs.) if moderately dehydrated	HW: - gives S-P - gives S-P in correct dosage - asks to wait 30 min. to check for vomiting - refers if signs of cerebral malaria	HW: - treats with antibiotic only when pus and ear pain - wicks ear - treats fever if present - refers case if swelling behind ear
3. PHYSICAL EXAM	HW: - advises on continued feeding during and after illness - advises on increasing fluid intake - advises to soothe throat and relieve cough - tells caretaker to return quickly if breathing becomes difficult - tells caretaker to return quickly if child is unable to drink - tells caretaker to return quickly if child becomes more ill - tells caretaker to return if child has convulsions - tells caretaker to return in 2 days for reassessment if treated	HW: - advises caretaker to continue feeding during diarrhoea - advises to give extra feeding after diarrhoea - advises caretaker to increase fluids until diarrhoea stops - tells caretaker how to prepare ORS - tells caretaker how to administer ORS - tells caretaker to return in 3 days if child does not improve - tells caretaker to return quickly if danger signs of dehydration appear - tells caretaker how to prevent diarrhoea	HW: - advises on continued feeding during and after illness - advises on increasing fluid intake - advises on sponging or bathing child with water to bring down fever - tells caretaker to return in 3 days if fever persists - explains that malaria can be prevented by bednets or screens - discusses how cleaning the environment prevents malaria	HW: - advises on wicking ear - tells caretaker to return in 5 days for reassessment if treated

were noted for each encounter, as well as the number of cases of child illness seen at the facility per day. Observations and interviews were conducted by Nurse Tutors, who participated significantly in the adaptation of the survey instruments to the Malawi context. The data collection forms used can be found in Appendix B.

The data entry and analysis were done using EpiInfo, SPSS/PC+, and Excel (for the cost data). The Kappa statistic,³ which tests agreement between two measures taken on the same sample, was used to assess agreement between the various quality assessment methods. A statistically significant value of Kappa is unlikely to be due to chance, and when significant, the size of Kappa indicates the level of agreement. Standards for rating agreement beyond chance, based on Kappa, are as follows:

0.00-0.39 = poor agreement [*values approaching 0.40 have been labeled: almost fair*]

0.40-0.74 = fair to good agreement

0.75-1.00 = excellent agreement

All Kappa and their P values were calculated using EpiInfo's epicalculator.⁴

For the comparison of provider interview results with observations, observation data were aggregated by provider and a composite score was developed in order to compare an individual provider's multiple patient encounters with his/her interview response. The composite score was based on the following criteria: if the provider was observed carrying out the task during at least one observed encounter, this was counted as a positive response when compared to the provider's response to questions about

⁴ Centers for Disease Control and the World Health Organization. *Epi Info: A Word Processing, Database, and Statistics Program for Public Health. Version 6 01*. Center for Disease Control, Atlanta, 1994.

³ Fleiss, JL. *Statistical Methods for Rates and Proportions*, 2nd Edition. John Wiley and Sons: New York, 1981.

† These are booklets prepared by the Ministry of Health and distributed to all health workers. They were developed with support from the Malawi Essential Drug Programme, and provide standards for history-taking, physical examination, diagnosis, treatment, counseling, and follow-up for common conditions.

whether he/she carried out that specific task.[§] The analysis was first done by individual task for each main symptom (cough, diarrhoea, fever, and ear problems), and then compiled into general tasks per main symptom, and finally across all main symptoms for a single child.

Consistency of provider performance on individual tasks was assessed by two measures: 1) the percentage of providers who either performed the task on all of their patients or none of the patients (% consistent), and 2) the percentage of providers who, if they were observed performing the task on at least a single patient, were also observed performing it on all their patients (% always doing if ever done). These analyses were performed only for providers who were observed for three or more patients.

IV. Results

A. *Ability of Quality Assessment Methods to Identify Children with Main Symptoms (Cough, Diarrhoea, Fever, and Ear Problems)*

The analysis compared the ability of three methods — observation, exit interview and record review — to classify individual children as having one of the four main symptoms: cough, diarrhoea, fever, and ear problems. A method's ability to identify all major symptoms is important in order to judge the correctness of the treatment provided.

Generally speaking, the agreement between the observation and exit interview data were excellent. In those cases of non-agreement, it was most often the exit interview that identified a case that was not so identified through observation (cough and fever). These extra cases included symptoms not spontaneously reported to the health worker or that the health worker did

§ It was not possible to devise a weighted score, since the number of cases seen by each provider of each main symptom were not the same. For example, in the case of diarrhoea, some providers were only observed seeing one case, while others saw as many as 15 cases.

not ask about. Agreement between observation and record review for those cases found in both was also excellent, although Kappa values were lower than those seen between exit interviews and observation. Consistently, non-agreement between record review and observation was due to cases identified by observation but not by record review; symptoms of cough and fever were most left out in the patient register, indicating that providers considered them associated symptoms of the major complaint. Rates of agreement between exit interview and record review were lower still, but generally were good to excellent. Again, more cases were identified by exit interview than by record review.

The percent agreement and Kappa values for the comparison of observation with exit interview, observation with record review and record review with exit interview are presented in Table A-1 in Appendix A. All the Kappa statistics have P values less than 0.0001, meaning that the level of agreement observed is not due to chance.

B. Comparison of Observation, Exit Interviews, Record Review and Provider Interview Data on Management of the Sick Child

The following sections present the comparisons of three assessment methods to observation, for the basic tasks of management of ill children: general assessment, management of cough, management of diarrhoea, and management of fever. Due to the small sample size ($N = 27$), management of ear problems has not been included.

The analyses were conducted based on cases where children were identified by both assessment methods. As a result, the numbers of children being compared for one pair of assessment methods are not the same for comparison of another pair of assessment methods. For example, comparisons of record review and observation only include 363 children for general assessment tasks, while comparisons of exit interview and observation include 423 children.

For Tables 2-6, Kappa values shown in bold are those that indicate fair or good agreement (greater than 0.400).

1. General Assessment Tasks for Assessment of the Sick Child

There are three major aspects to the general assessment of the sick child: determining the presence of danger signs, obtaining information about presence of main symptoms, and providing preventive services, such as immunizations and growth monitoring. Table 2 provides Kappa values comparing exit interview, record review, and provider interview data to observation of these tasks. (See Table A-2 in Appendix A for percent agreement and prevalence of task performance detected by each method.)

Table 2			
General Assessment of the Child			
Task	Kappa Statistic for Method Comparison		
	OBS-Exit Interview N=423	Obs-Record Review N=363	OBS-Provider Interview N=30
Ask about ability to drink	0.189*	0.00	0.053
Ask about convulsions	0.381*	0.00	-0.101
Assess consciousness	-0.284	0.00	-0.027
Ask presence of cough	0.634*	0.459*	**
Ask presence of diarrhoea	0.547*	0.484*	**
Ask presence of fever	0.526*	0.467*	**
Ask presence of ear problems	0.329*	0.311*	**
Check immunization status	0.134*	0.00	-0.056*
Assess growth	0.231*	0.0003	-0.065*
Weigh child	0.684*	0.00	-0.174*
Plot weight	0.453*	0.00	0.00

* P < 0.001

**Item not included in data collection instrument

These data indicate that record reviews were only able to furnish general assessment information about presence of main symptoms, and occasionally nutritional status, and that provider interviews did not provide reliable information about what workers do. Therefore, the following discussions will focus on comparison of exit interview and observation data.

Danger Signs: Table 2 indicates that information about looking for danger signs can most reliably be done through direct observation of the patient-provider encounter. Mothers do not appear to be able to recognize actions that a health worker may carry out to assess the state of consciousness, and exit interviews showed ‘almost fair’ agreement only for questions about convulsions related to this illness.

Important Symptoms: Both exit interview and record review showed fair to good agreement with observation data about whether the health worker obtained information about the presence of individual main symptoms (cough, diarrhoea, fever, and ear problems). For observation data, this would be that either the health worker asked specifically, or the mother volunteered the information. For exit interviews, these rates reflect mothers reporting that health workers asked specifically. For record review, these data reflect whether the record indicates any symptoms or disease associated with that symptom (i.e., pneumonia would indicate cough, and malaria would indicate fever). Agreement with observation was stronger for exit interviews than for record reviews.

Preventive Services to the Sick Child: Comparisons with observation for preventive health care activities (checking immunization status, assessing growth, weighing the child, and plotting the child’s weight) indicate that exit interviews can detect concrete activities, such as weighing and plotting of weight on the under five card, but only observations appear able to reliably identify more refined tasks of checking immunization status and growth: mothers do not appear able to interpret what tasks a provider is performing while examining the Under Five card.

2. Specific Tasks related to Individual Main Symptoms

The following sections compare exit interview, record review and provider interview data with observation data related to specific history, physical, treatment, and counseling tasks.

A. Assessment, Treatment, and Counseling of Children with Cough

Table 3 presents the results of comparison for the individual tasks in managing a child with cough. As with general assessment tasks, record review was able to provide limited information, only showing excellent agreement for treatment, and provider interviews did not generally have much agreement with observation data. Since exit interview data displayed fair to excellent agreement on many tasks, presentation of results will focus on comparisons of exit interview with observation data. (For complete data, see Table A-3 in Appendix A.)

History: Exit interview data showed ‘almost fair’ and fair agreement with observation data for questioning related to cough. Provider interviews also had ‘almost fair’ agreement for questions about difficult breathing.

Physical Exam: The only examination task that mothers were able to reliably report on was observing the child for chest indrawing (expose child’s chest and watch him breathe).

Diagnosis and Treatment: Agreement about classification of severity of illness with data from observation was poor for all methods; however, it is also likely that the observation data are also not very reliable, without the aid of asking questions.** Agreement on whether the provider told the mother a diagnosis of her child’s illness was also very poor, with mothers rarely stating they had been told their child’s problem. Agreement on treatment was excellent for both exit interview (taken from the OPD ticket) and record review (taken from the registers).^{††}

Counseling: Mothers appear to be able to fairly reliably report on counseling given about feeding during illness, returning if breathing becomes difficult or child becomes more ill. Record review provided information related to counseling only when treatment was continued breastfeeding and sips of water.

Table 3			
Case Management of Cough			
Task	Kappa Statistic for Method Comparison		
	OBS-Exit Interview N=423	Obs-Record Review N=363	OBS-Provider Interview N=30
Ask about duration of cough	0.464*	0.009	0.132
Ask about difficult breathing	0.347*	0.026	0.318+
Count respiration rate	0.231*	0.000	-0.085
Observe for chest indrawing	0.550*	0.000	0.154
Check for neck stiffness	-0.031*	0.000	-0.098
Classify child by severity of illness	0.027	0.129	0.231
Tell diagnosis of cough	0.114*	**	**
Give appropriate "drug" for cough	0.755*	0.780*	**
Give correct cough RX dosages	0.718*	0.754	**
Tell to feed during illness	0.416*	0.27	0.49
Tell to increase fluid intake	0.369*	-0.069	0.133
Advise to soothe and relieve cough	0.238	0.112	0.153
Tell to return if breathing is difficult	0.496*	0.000	0.121
Tell to return if child is unable to drink	0.093	0.000	0.067
Tell to return if child is more ill	0.436*	0.000	0.253+

* P value is < 0.001

+ P value is < 0.05

** Item not included in data collection instrument

B. Assessment, Treatment, and Counseling of Children with Diarrhoea

Table 4 presents the results of comparison for the individual tasks in managing a child with diarrhoea. (See also Table A-4 in Appendix A). Again, record review furnished limited information, only on treatment. However, both exit interview and provider interview showed better agreement for diarrhoea tasks than that seen for cough. Presentation of results concentrates on comparisons of exit interview and provider interview with observation data.

History: Information from exit interview displayed fair to good agreement with observation data, with the exception of whether the child was urinating.

Physical Exam: Mothers were less able to report reliably on physical examination tasks, except for checking for a sunken fontanel, which showed 'almost fair' agreement.

Diagnosis and Treatment: No method was able to provide reliable data on the diagnostic process. Mothers were not able to reliably tell whether they had been told the diagnosis (percent agreement of 49%). For treatment, correspondence was excellent for both record review and exit interview, because both these sources used the outpatient ticket. Correspondence of provider interview with observation data also showed 'almost fair' agreement for treatment of dehydration.

Counseling: Mothers were reliably able to report advice about feeding and fluids (agreement fair to good) and about when to return (agreement fair), but agreement was poor for advice on how to prepare and administer ORS and for prevention. Provider interview showed fair agreement on feeding, fluids and preparation of ORS, but poor agreement for advice on when to return.

Table 4			
Case Management of Diarrhoea			
Task	Kappa Statistic for Method Comparison		
	OBS-Exit Interview N=123	OBS-Record Review N=86	OBS-Provider Interview N=29
Ask about duration of diarrhoea	0.567*	0.005	0.298+
Ask about blood in stools	0.477*	0.127	0.188
Ask whether urinating	0.056	0.000	-0.105
Pinch skin	0.241*		0.220+
Check for sunken eyes	0.228*	0.000	-0.080
Check fontanel (N = 61)	0.390*	0.000	0.083
Classify degree of dehydration	0.026	0.263	-0.125
Treat dehydration appropriately	-0.080	0.183	0.337
Tell diagnosis of diarrhoea	0.133	**	**
Give ORS	0.617*	0.178	**
Diagnose normal diarrhoea	0.922*	0.945*	**
Give appropriate "drug" for diarrhoea	0.906*	0.958*	**
Give correct diarrhoea RX dosages	0.843*	0.958*	**
Tell to feed during illness	0.661*	0.000	0.429+
Tell to feed after illness	0.415*	0.000	0.319+
Tell to increase fluid intake	0.564*	0.040	0.602+
Explain how to prepare ORS	0.286*	0.000	0.343+
Explain how to administer ORS	0.298*	0.000	0.376+
Tell to return if signs of dehydration	0.393*	0.000	0.096

* P value is < 0.001

+ P value is < 0.05

** Item not included in data collection instrument

C. Assessment, Treatment, and Counseling of Children with Fever

Table 5 presents the results of comparison for the individual tasks in managing a child with fever. (See also Table A-5 in Appendix A.) In summary, record review provided information only on treatment, and provider interview did not generally show much agreement with observation data. Exit interview showed better agreement for physical exam tasks than for cough or diarrhoea, and presentation of results focused on comparisons of exit interview with observation data.

History: Exit interviews with mothers showed fair to good agreement for questions specifically related to fever. However, mothers appear less able to report accurately on whether providers asked about other symptoms that could explain the presence of the fever.

Physical Exam: Mothers were quite able to report with good to excellent agreement for all physical examination tasks, with the exception of checking for anemia. Although the investigators had expected that mothers would not notice providers checking for splenomegaly (press on child's abdomen on upper left side), this showed fair agreement.

Diagnosis and Treatment: As with cough and diarrhoea, mothers tended not to report that the provider told them a diagnosis for their child. Agreement on treatment given was excellent for both exit interview and record review, and although provider interview data showed low Kappa[‡], the percent agreement values are high. Agreement about whether providers tell mothers to wait 30 minutes after their child has taken the malaria treatment was 'almost fair' for both exit interview and provider interview.

Counseling: Mothers appear able to report reliably on receiving counseling messages on feeding during illness and prevention of malaria. Other messages show 'almost fair' agreement.

[‡] When the proportions are very large, Kappa is low as the chance of not agreeing would be small anyway.

Table 5			
Comparison of Observation, Exit Interview, Record Review, and Provider for Interview for Fever/Malaria			
Task	Kappa Statistic for Method Comparison		
	OBS-Exit Interview N=200	Obs-Record Review N=160	OBS-Provider Interview N=29
Ask about presence of other symptoms	0.111+	0.066+	0.125
Ask about vomiting	0.400*	0.000	0.058
Ask about treatment already taken	0.4901*	-0.013	0.126
Take temperature	0.796*	0.000	0.015
Check for splenomegaly	0.410*	0.000	0.004
Take blood for a slide	0.496*	0.000	-0.063
Check for anemia	0.384*	0.000	0.231
Give S-P	0.867*	0.860*	**
Ask to wait for 30 minutes	0.374*	0.000	0.312+
Tell diagnosis of malaria	0.180*	**	**
Give appropriate "drug" for malaria	0.683*	0.813*	-0.046
Give correct malaria RX dosages	0.614*	0.831*	-0.060
Tell to feed during illness	0.432*	0.000	0.267
Tell to feed after illness	0.391*	0.000	0.254
Tell to increase fluid intake	0.380*	0.133*	0.117
Advise to give child sponge bath	0.371*	0.081+	-0.114
Tell to return in 3 days if not better	0.309*	0.000	0.098

* P value is < 0.001

+ P value is < 0.05

** Item not included in data collection instrument

D.General Comments about Information from the Various Quality Assessment Methods

Lack of agreement between provider or mothers' interviews and observation can be due to either under-reporting or over-reporting. Differences in rates reported from various sources could lead to different conclusions about provider performance. A general review of levels reported from exit interview and provider interview in comparison with observation data showed the following tendencies:

Exit Interview Data:

- **History:** Mothers showed no patterns—for some tasks reporting more, for some reporting less.
- **Physical Exam:** Patterns here depended on the specific health problem. For general assessment and diarrhoea tasks, mothers tended to under-report. For fever tasks, the levels were generally about the same (and agreement was good), while cough tasks tended to be slightly over-reported by mothers.
- **Treatment:** For fever and diarrhoea, data obtained during the exit interview from the OPD ticket provided similar rates of correct drugs and treatment schedules as observation data, but for cough, exit interview data tended to slightly under-reported.
- **Counseling:** For fever, mothers tended to report similar levels, but for diarrhoea and cough, mothers tended to over-report both for messages on home care and on when to return.

Provider Interview Data:

- **History:** Providers tended to under-report history questions they ask for both diarrhoea and fever, while over-reporting for cough and general assessment.
- **Physical Exam:** Providers tended to over-report physical examinations for all health problems, with the exception of general assessment.

-
- **Treatment:** Data for fever showed providers reporting what they were actually observed doing, while data for diarrhoea indicate providers over-reporting correct treatment of dehydration. No data were available for comparison for cough.
 - **Counseling:** For counseling messages related to home care, providers reported rates similar to those observed. However, for messages relating to when mothers should return, providers tended to over-report.

3. Integrated Management of the Sick Child

Of the 436 child patient-provider encounters observed, 32% were cases that had more than a single main symptom/health problem. Seventy-five children (17%) presented with cough and fever, 54 (12%) presented with fever and diarrhoea, and 20 (5%) with cough and diarrhoea. Eleven children (3%) presented with cough, fever, and diarrhoea. Cases of cough and fever, and of fever and diarrhoea were examined for provider performance with respect to the two symptoms together. The analysis was carried out by looking at the number of encounters where at least one history question was asked for each symptom, at least two history questions were asked for each symptom, etc. This type of analysis provides a measure of integrated management since it requires that the health worker exclude through the assessment process that there are not serious complications related to either symptom (i.e., whether a case of cough+fever is really a respiratory infection or malaria with a cold). Treatment was assessed based on treatment of both conditions (e.g., fever with either aspirin or S-P; diarrhoea always receiving ORS, even if associated with malaria; cough not receiving aspirin except in the presence of fever). Integrated counseling would involve providing the appropriate messages for all symptoms (about when to return and about what to do at home).

Table 6 presents the comparison of observation and exit interview data on integrated management of fever+diarrhoea and cough+fever (see also Tables A-6 and A-7 in Appendix A). Comparisons with record review and provider interview were not included, as record review did not provide information on most tasks, and provider interview were not structured to address integrated management of ill children.

The data in Table 6 indicate that exit interviews with mothers provide reliable data on integrated management of fever+diarrhoea. Agreement between exit interview and observation was good for all aspects of case management except counseling on when to return. Reliability was not as good overall for cough+fever, since agreement was not high for history questions and for advice on when to return. It should be noted that although this analysis used generic indicators, it was based on very specific data which were combined into generic indicators in the analysis.

Table 6		
Integration Management of the Sick Child		
Task	Kappa Statistic for Comparison of Observation and Exit Interview	
	Fever and Diarrhoea N=54	Cough and Fever N=75
Asked 1 or more history question for each problem	0.412*	0.293*
Asked 2 or more history questions for each problem	0.547*	0.374*
Asked all history questions for each problem	–	-0.013
Did 1 or more physical exams for each problem	0.441*	0.421*
Did 2 or more physical exams for each problem	0.274+	0.211+
Did all physical exams for each problem	–	–
Treated correctly for both problems	0.476*	0.546*
Gave any advice on feeding	0.522*	0.419*
Gave all advice on feeding	0.457*	–
Gave advice on home care for each problem	–	–
Gave any advice on when to return for each problem	0.315+	0.056

* P value is < 0.001

+ P value is < 0.05

C. *Consistency of Provider Behavior*

One of the critical issues in obtaining information about provider behavior is whether providers actually perform in a consistent manner with all patients. If they do perform consistently, it would only be necessary to observe them managing a single child or interviewing a single mother to have information about a provider's routine performance. To examine the level of consistency of provider behavior, two measures of consistency were developed: 1) percentage of providers who carried out a single task either on all their patients or none of their patients (% consistent), and 2) percentage of providers who, if they were observed performing on a single patient, were also observed performing it on all their patients, with that main symptom (% doing if ever done). These measures both use the provider as the unit of analysis.

Figures 1-3 present these results for cough, diarrhoea, and fever patients. (For additional data, see Tables A-8, A-9 and A-10, respectively, in Appendix A). It should be noted that the number of provider-patient encounters observed for each provider covered a wide range. For cough, the number of cases ranged from 1 to 21, with an average of 6.73 and a median of 4. For diarrhoea, the numbers ranged from 1 to 15, with an average of 4.31 and median of 3. For fever, the range was 1 to 17, with a mean of 6.91 and a median of 5. For the data presented in Figures 1-3, only providers who were observed managing 3 or more cases were included in the analysis.

Management of Children with Cough: A total of 30 providers were observed managing 3 or more children with cough. As seen in Figure 1, proper provider behavior, as measured by following the prescribed standard tasks for history, physical, treatment and counseling, was generally inconsistent. Less than two thirds of providers showed consistent behavior for most of the tasks. Where this consistency measure shows higher values, it is because providers were consistent in never performing those tasks (e.g., counting respiration rate, checking for neck stiffness, telling mother to return if child is unable to drink or has convulsions). Inconsistency is also reflected in the measure % always doing if ever done: less than one-fifth of the providers who ever carried out a specific cough management task did it every time. The only exceptions were asking about duration of cough and telling the mother a diagnosis, where nearly half the providers did so for every patient.

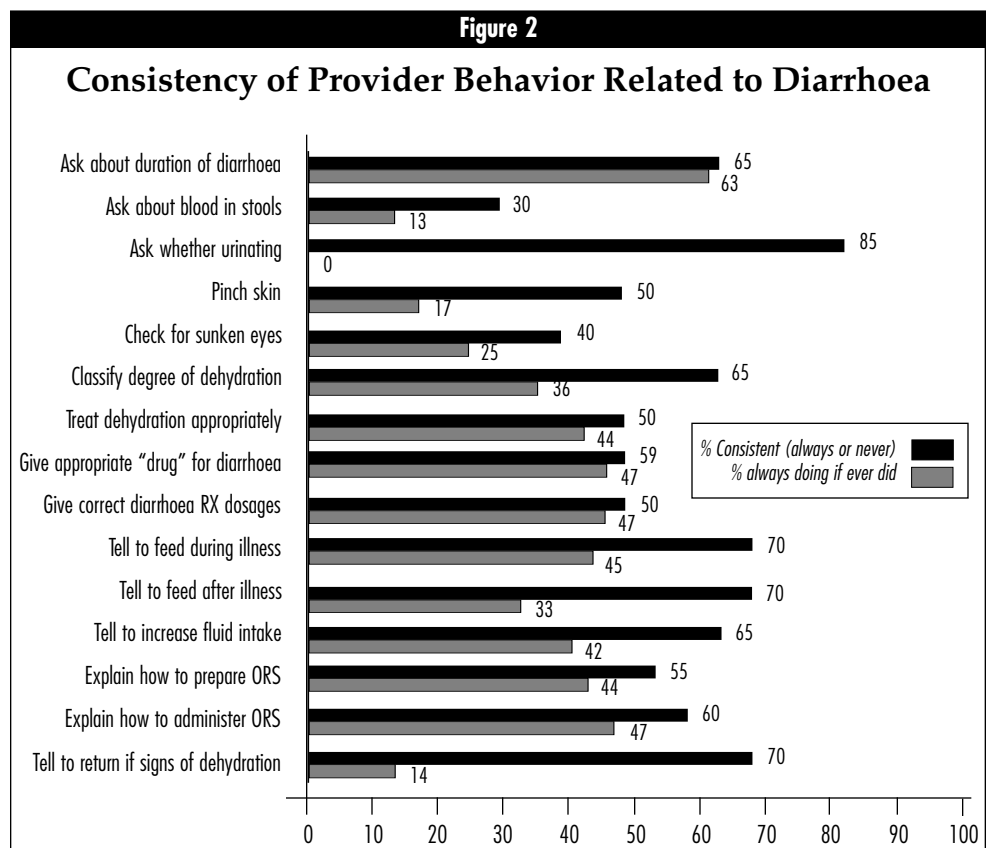
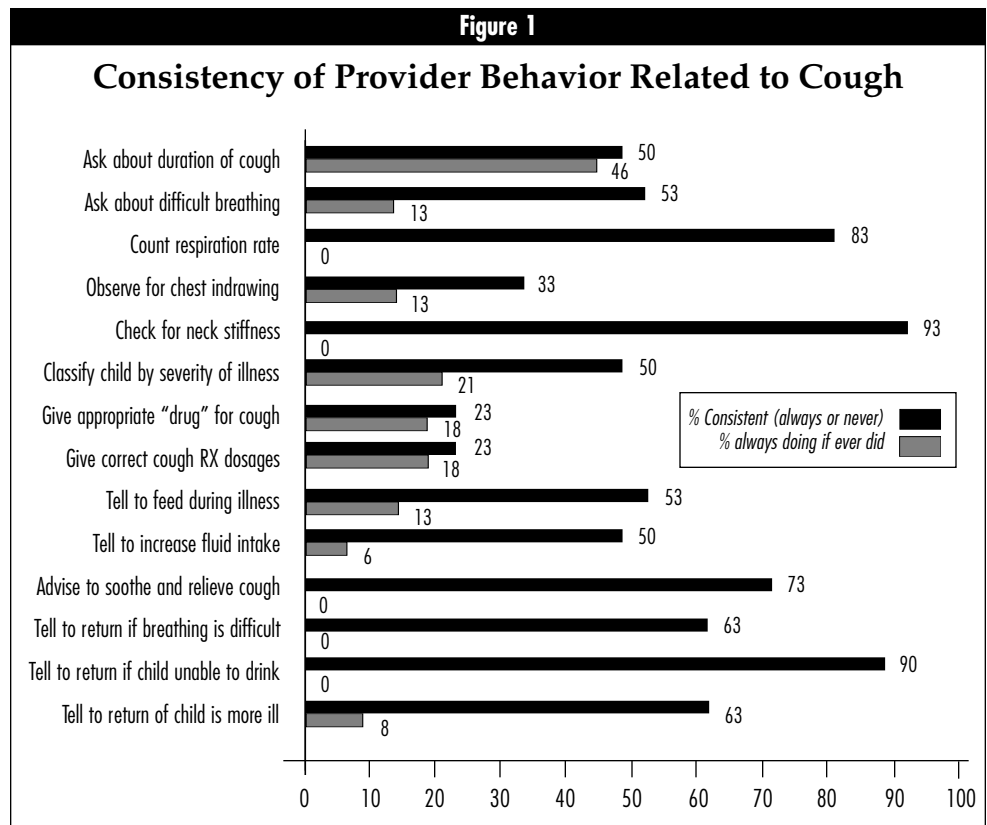
Management of Children with Diarrhoea: Figure 2 shows the specific consistency data for diarrhoea case management. As with cough, providers tended towards inconsistent behavior, and where the % consistent measure was higher, these were often for tasks that were rarely performed by providers. However, a higher percentage of providers were seen performing the tasks on all their patients if they were observed ever doing it: most tasks show a third to half of providers that ever did the task performing it on all patients observed. This is particularly true for counseling tasks and appropriate treatment.

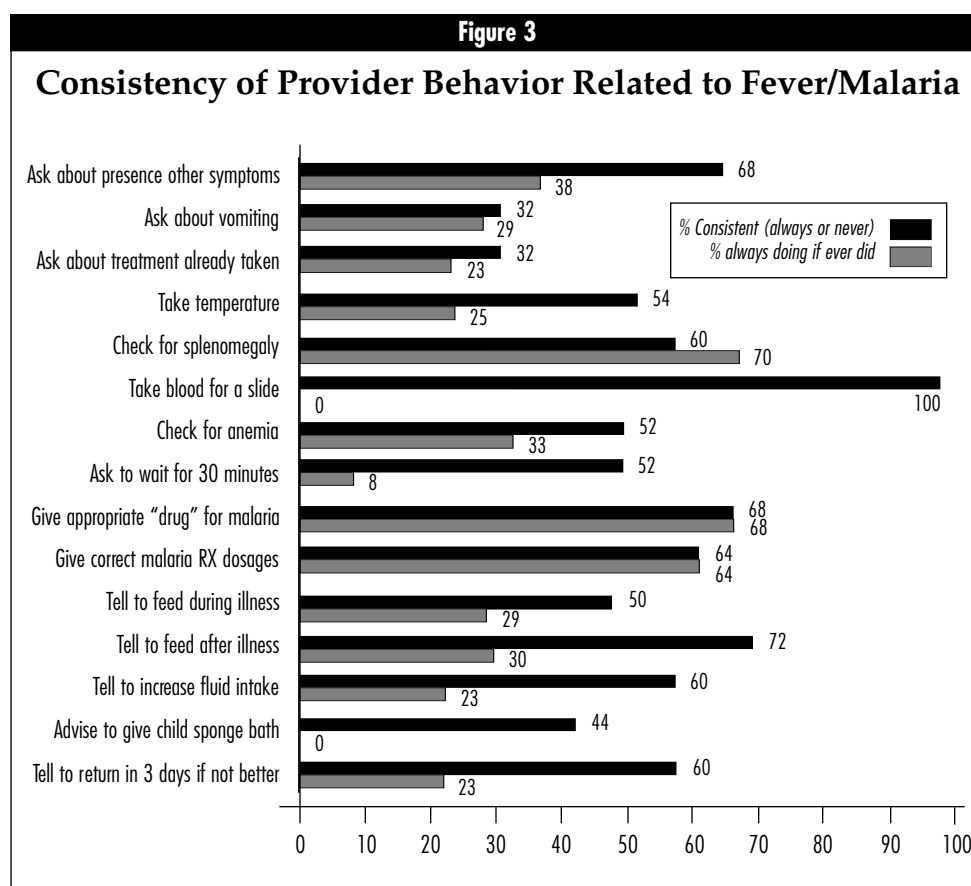
Management of Children with Fever: Figure 3 shows these same measures for management of fever patients. Again, where % consistent measures were high, it was usually because the providers never performed the task. The percentages of providers consistent if they were ever observed performing the task were not as high as for diarrhoea, but were higher than for cough. Percentages for correct treatment were the highest.

These data were also examined to see whether patient order played a role in provider inconsistency, due to either initial unease in the presence of observers or due to a return to normal behavior as one began to “forget” the presence of the observers. No association was found between patient order and provider performance.

The data do not indicate how many cases a supervisor would need to observe to get a reliable picture of provider performance, but they do suggest that single observations are not sufficient. It is worth noting that providers were more consistent in diarrhoea case management than for fever or cough. This could be related to health worker training and refresher training, which has been strong in the area of diarrhoea management.

There is an additional aspect to provider consistency that relates to the ‘art of medicine.’ Providers may look at a child and from his state, determine whether this child is really very ill or not. Based on that assessment, they may forego certain questions or exams that are not really indicated in a basically well child. The different possible states of the child are not well accommodated in a standardized checklist.





D. *Cost of Various Methods of Quality Assessment*

The resources required to use any of these quality assessment methods for supervisory purposes include the supervisor's time, travel costs, and costs of supplies (forms, etc.). However, all these methods require the supervisor or an interviewer to be present at the health care facility. Thus, the differentiating factor in cost is the supervisor's time. The resources needed to employ the four assessment methods were thus calculated in several ways: the average time required to assess a single case of any illness, the time required for exit interview and observation to see at least one case (number of cases per hour of observation or interviewing), and the time required to see at least one serious case.

1. *Time Required to Use the Methods*

The times required by each method to assess case management of a single ill child (or interview of a single provider) can be seen in Table 7.

Table 7			
Time Required to Assess Quality per Case (III Child or Provider) (in minutes)			
Method	Mean	Median	Mode
Observation (N=432)	3.48	2.5	2
Exit Interview (N=420)	6.60	5.5	5
Record Review (N=348)	3.03	2.5	2
Provider Interview (N=29)	34.79	34.5	30

The time for conducting observations of case management of ill children reflects the time the provider spent with the patient. However, the actual observation form took longer to fill in (anecdotally, about an additional 2 minutes), and observers were not able to record information for all children presenting with cough, fever, diarrhoea or ear problems. In these cases, observers skipped the next child, even if the child had one or more of the main symptoms, to allow themselves time to fill in their forms completely for the one just observed.

The results from this table indicate that extracting information from a simple patient record takes about the same time as the actual observation, although if the data to be collected were limited to that normally found in the patient record (treatment and referral), it would take less time to conduct a single record review. Although the provider interview takes longer than the other methods, this is not directly comparable to other methods, since one would conduct a single interview, while the other methods would be done on more than one child (see discussion on consistency of provider behavior in Section IV.C.).

2. Time Required to See a Single Case of Each Main Symptom

One of the difficulties of observation and exit interview is that they are dependent on the volume of patients seen during the observation or exit interview period. This means that even though health centers saw an average of 34 ill children under five in a single day, it does not mean all these children had the illnesses of interest. The average number of ill children per hour of observa-

tion was 6.4, and the average number of hours of operation of the OPDs was 5.37 hours, with ranges from 3.3 hours to 6.42 hours (not including lunch hour).

Table 8 show the average number of cases of each health problem seen in a one-hour period. It should be noted that these data were collected in the cold/dry season when OPD attendance is not at its peak, and although there are a greater number of respiratory problems, diarrhoea and malaria are not at their most prevalent. These data average over the entire duration of operation per day; however, cases tend to be clustered in the morning hours.

Table 8			
Number of Ill Children per Hour during the Cold/Dry Season N = 40 days of observation			
Method	Mean	Median	Mode
Cough	1.39	1.2	1.1
Diarrhoea	0.82	0.6	0
Fever/malaria	1.18	1.1	0.9
Ear Problem	0.17	0	0

3. Time Required to Assess Management of Serious Cases

One critical aspect of assessing quality of provider management of sick children is their ability to manage the very ill child. However, cases of pneumonia, cerebral malaria, and moderate or severe dehydration are generally rare. Two criteria for determining the number of very ill children seen in the Under Five Clinics are presented here: 1) provider's diagnosis from observation data, and 2) mother's reporting of specific symptoms from the exit interview.

Cough: Based on the provider's diagnosis of pneumonia, severe pneumonia and very severe pneumonia, only 11% of the 217 children observed with cough were so diagnosed. Using the mothers' reporting of child having difficulty breathing, 18% of the 235 children whose mothers said they had cough also reported difficult breathing.

Diarrhoea: Using provider's diagnosis of the degree of dehydration, 6% of the 124 children observed with diarrhoea had either moderate or severe dehydration. Asking the 132 mothers who reported their children having diarrhoea, 8% said the provider had told them their child had lost lots of fluids.

Fever/Malaria: No cases of cerebral malaria were diagnosed by providers or reported by mothers.

These data indicate that getting information about management of very ill children through observation or exit interview is difficult under normal dry/cold season conditions.

E. Evaluation of Quality Assessment Methods: Strengths and Weaknesses for Obtaining Data on Quality

Each quality assessment method has its strengths and weaknesses. Some of these are inherent in the methods themselves, and some are related to their ability to provide valid or reliable data about certain specific tasks. Each method will be discussed separately.

1. Observation of Provider Performance Using a Checklist

As mentioned in Section IV.C., most providers are generally not very consistent in their performance, and thus observation data must include multiple provider-patient encounters in order to say something reliable about what the provider typically does. Observation data are also sometimes difficult to collect, since management of an individual sick child may be performed by several staff in different locations: consultation room, injection room, pharmacy, ORT corner. In addition, providers often see patients for only a few moments, meaning observers do not have time to finish filling in the checklist before the next patient is being seen.

Observation data also have certain constraints related to their ability to record the mental steps that providers take during their diagnostic process and even some aspects of the physical examination. These difficulties were reported by observers for classification tasks and for the general assessment of the child related to state of consciousness. Observation checklists also ask observers to

follow a prescribed order to the tasks, while providers may not follow that order at all. This makes it hard for observers to complete the checklist correctly. Inter-observer reliability (two individuals observing the same provider-patient encounter) conducted during training showed that it was possible to miss individual tasks.

Additional problems arose for observation of integrated management of the sick child that would not necessarily be problematic if only management of single main was being assessed. It was difficult to determine which history, physical examination and counseling tasks were appropriate tasks, when it was unclear whether fever or diarrhoea were to be considered as a main or an associated symptoms. Using the example of a child with diarrhoea and fever: is this a case of fever being associated with a bowel infection or the case of malaria with associated diarrhoea? This same issue arose when trying to rate correctness of treatments: should cough plus fever be considered essentially a respiratory problem, or malaria with an associated cold? Should diarrhoea plus fever be treated as gastric malaria or as a bowel infection?

However, despite these problems, data from direct observation allow the most complete information and can be collected by those trained to recognize specific tasks well. Because observers have a checklist, they are less likely to forget what happened in the encounter than would the patient or caretaker.

2. Exit Interviews with Caretakers about Provider Performance

As the comparison data in Tables 2-5 showed, mothers were often able to reliably report (good agreement with observation data) on what happened in the encounter with the provider. However, there were some areas that were particularly difficult: as with observation data, these include mental tasks of the provider, such as classification and assessment of state of consciousness. But mothers also had difficulty identifying which part of the Under Five card the provider examined (i.e., the side with immunization information, or the growth chart itself), certain parts of the physical examination, and whether any diagnosis was told to them.

The ability of exit interview to provide reliable data also depends on the memory of the mother, how much she was paying attention to what the provider was doing and her expectations about what the provider should be doing. For example, the low level of agreement about whether the provider

told the mother a diagnosis indicates that mothers may have different expectations about what constitutes a 'diagnosis' and were feeling that the providers had not said enough.

Exit interview has several advantages, however. It can be performed without the knowledge of the provider and therefore will not have the problem of the 'Hawthorne' effect where people change their performance because they know they are being observed. In addition, an exit interview provides the interviewer (if he has clinical training) with the opportunity to examine the child himself to get information about the accuracy of the diagnosis (although this was not done in this study). Such examination would, however, add to time required per interview.

3. Reviewing Records to Assess Provider Performance

The type of records kept in out-patient departments in Malawi (and in many other African countries) do not provide sufficient data to be able to assess most of the case management process. They do provide information about treatment based on the provider's diagnosis, but do not assist in determining whether the diagnosis is accurate or not. The data available (treatments) showed excellent agreement for all health problems, even though a variety of individuals were filling in the patient register, including a medical assistant, a nurse, a hospital servant, a ward attendant, a health surveillance assistant, and a ground laborer.

Record reviews do have the advantage of allowing perusal of numerous cases in order to assess treatment of diagnosed severely ill children. Although about 18% of patients observed were not found in the registers, this should not affect their ability to assess quality, as long as there is not a pattern among those omitted.

4. Interviews of Providers about their Performance

Provider interviews provide an opportunity to assess what providers know. Comparison data indicate that providers do not necessarily perform tasks that they report they do; this severely limits the accuracy of the provider interview method for gauging health workers' routine performance. However, interviews do provide data about knowledge of management of severely ill children which are not readily available from observation and exit interview and are not contained in records (accuracy of diagnosis).

There are two additional areas where provider interviews present difficulties:

- Probed versus spontaneous responses: Providers are not always able to remember all they do, when discussing verbally with someone about what they do when physically confronted with a patient. Only counting spontaneous responses is likely to be an underestimate of what they do, especially for tasks that they do not carry out often or for tasks that they do so often that it comes unconscious. Yet, probing to stimulate memory can also stimulate responses that are what the provider thinks the interviewer is looking for.
- Looking at integrated management: There are two possible ways to interview about case management: using a generic framework (regardless of main symptoms) or following the tasks for specific main symptoms. Neither of these are well suited to assessment of integrated management of children with multiple symptoms, which would require either case studies or simulations.

V. Comparison with Data from Other Studies and Conclusions

A. Comparison with Results from Other Studies

Only two similar studies in developing countries have analyzed data to compare reliability and validity of quality data from different sources: the sexually transmitted diseases (STD) health facility survey conducted in Malawi in 1994, and a Quality Assurance Project (QAP) methodology refinement study conducted in Guatemala, also in 1994.

1. Malawi STD Health Facility Survey

A health facility survey based on the WHO protocol for STD case management was conducted in Malawi in 1994, using the following quality assessment methods: direct observation, provider interview and simulated patients (posing as urethral discharge patients). A total of 49 providers were observed

and interviewed, and 20 of these were also visited by a simulated patient. Provider interview data were analyzed separately into spontaneous responses and probed plus spontaneous responses (as in this study). Comparative analysis was conducted using the Kappa statistic. Provider interview data showed large variations between probed and spontaneous responses, with spontaneous answers often being lower than observed performance, while probed responses almost always indicated better performance than that observed. In general, agreement between direct observation and provider interview was poor, with the exception of some aspects of the physical examination and treatment. Agreement between direct observation and simulated patients was also generally poor, regardless of whether direct observation was measured on the basis of performing the task on all patients or only on some patients.

2. Study of Comparative Validity of Several Methods for Assessing Health Worker Performance in Guatemala

A study similar in design to this one was conducted in Guatemala in 1994, comparing observation, mothers' interview and record review with a trained observer's observation. Nine providers were observed in 3 health centers. A total of 74 cases of acute respiratory infections, 56 diarrhoea cases and 67 family planning encounters were observed. A limited number of tasks for management of ARI, diarrhoea and family planning were assessed. Results indicate that interviews with mostly uneducated mothers provided reliable data for many tasks, with the exception of counseling on breastfeeding and asking whether the child has cough, where mothers over-reported. Records reviewed were generally not complete enough to provide information beyond weighing, temperature, and treatment. It should be noted that records in Guatemala are more complete than those in Malawi.

B. Conclusions

The results of this study, in conjunction with those from the two studies presented above, permit the following conclusions:

- For many case management tasks, exit interviews with mothers can provide reliable data about what happened in the provider-patient encounter.

-
- Provider interviews generally do not provide reliable data about what providers do.
 - Record reviews furnish only limited data for assessing case management.
 - Providers are not consistent in their management of individual patients.
 - The costs of observation and exit interview are similar in time, while record reviews take less time and are more easily scheduled. Provider interviews are the least expensive.

The tools used during this study could be improved in the following manner:

1. Discussions with providers during observation would enhance data on classification and diagnosis.
2. Examination of children during exit interviews would provide data on whether the provider had made a correct diagnosis, as well as aiding the mother to recognize certain physical examination tasks.
3. Adding case studies to the provider interview would allow assessment of integrated management of an ill child with multiple main symptoms.

C. Recommendations

From the data presented in this study, it appears that a hierarchy of methods should be applied by supervisors. This hierarchy follows the necessary pre-conditions for routine correct performance: knowledge, skills, application (practice).

Assessing Knowledge: If supervisors are unfamiliar with the knowledge levels of their supervisees, quality assessment should commence with a combination of: 1) provider interviews to assess whether providers have the knowledge necessary to conduct proper case management and 2) record reviews to assess whether they use the correct drugs and dosages for the diagnosis they ascribe. These methods are the easiest to administer and can even be done at a central location if transport for supervision is a constraint. If knowledge is found lacking, supervisors could either provide on-the-job training or arrange for formal training.

Assessing Skills: Once supervisors are confident in their supervisees' knowledge, it is time to assess their skills, through direct observation of performance. By using these occasions to discuss with supervisees, supervisors can correct deficiencies in performance. Supervisors would need to spend enough time with providers to observe multiple cases being managed.

Assessing Routine Application: When supervisors feel confident about their supervisees' ability to correctly manage cases in their presence, occasional assessment using exit interviews with mothers without the knowledge of the supervisee would allow supervisors to know what is routinely being done. Problems in case management arising at this stage would most probably be due to organizational or motivational causes, since knowledge and skill deficiencies would have already been corrected.

Finally, there are several areas requiring further research. One would be to determine the minimum number of observations/exit interviews needed to make a reliable judgment about provider performance. Another area would be improving the provider interview to accommodate more readily assessment of integrated management of the ill child. Finally, use of detailed checklists (specifying all the tasks for each symptom) need to be compared with generic checklists in terms of their ability to identify problems in integrated case management.

Appendix A

Table A-1				
Comparison of Methods for Identification of Children with Main Symptoms (% of children having syndrome)				
Observation-Exit Interview N=426 Children				
	% Observation	% Exit	% Agreement	Kappa
Cough	51%	55%	92%	0.840*
Diarrhoea	29%	31%	98%	0.944*
Fever	51%	51%	92%	0.836*
Ear Problems	6%	7%	99%	0.918*
Observation-Record Review N=363 children				
	% Observation	% Exit	% Agreement	Kappa
Cough	52%	43%	88%	0.764*
Diarrhoea	29%	25%	93%	0.826*
Fever	50%	44%	90%	0.807*
Ear Problems	6%	5%	99%	0.871
Record Review-Exit Interview N = 352 children				
	% Observation	% Exit	% Agreement	Kappa
Cough	43%	57%	84%	0.688*
Diarrhoea	26%	30%	94%	0.838*
Fever	44%	51%	85%	0.699*
Ear Problems	5%	7%	98%	0.810*

* P value is < 0.0001

N.B. Percentages presented are not cumulative; 32% of children had more than one syndrome

Table A-2												
General Assessment of the Child												
Tasks to be carried out	Obs-Exit Interview N=423				Obs-Record Review N=363				Obs-Provider Interview (1) N=3			
	%obs	% exit	Agree	Kappa	%obs	% rec	Agree	Kappa	% obs (3)	%	Agree	Kappa
Ask about ability to drink	9%	31%	72%	0.189*	9%	0%	91%	0.00	27%	33%	60%	0.053
Ask about convulsions	5%	11%	91%	0.381*	4%	0%	96%	0.00	27%	47%	47%	-0.101
Assess consciousness	34%	31%	43%	-0.284	33%	0%	67%	0.00	34%	63%	45%	-0.027
Refer if danger signs	1%	1%	numbers too small		1%	1%	numbers too small		no data available			
Give first dose before referral	1%	1%	numbers too small		1%	1%	numbers too small		no data available			
Ask presence of cough (2)	68%	60%	83%	0.634*	68%	43%	72%	0.459*	no data available			
Ask presence of diarrhoea (2)	51%	45%	77%	0.547*	49%	26%	74%	0.484*	no data available			
Ask presence of fever (2)	68%	70%	80%	0.526*	68%	45%	72%	0.467*	no data available			
Ask presence of ear problems (2)	23%	15%	79%	0.329*	22%	5%	83%	0.31	no data available			
Check immunization status (4)	95%	89%	87%	0.134*	95%	0%	5%	0.00	97%	87%	83%	-0.056
Assess growth (4)	94%	96%	95%	0.231*	95%	3%	6%	0.0003	97%	63%	60%	-0.065
Weigh Child	74%	65%	87%	0.684*	75%	0%	25%	0.00	87%	83%	70%	-0.174
Plot weight (4)	97%	96%	97%	0.453*	96%	0%	4%	0.00	100%	67%	70%	0.00

(1) In provider interviews, a probed YES was counted as NO (only spontaneous answers counted)

(2) For record review, these refer to records that include any mention of symptoms or diseases related to symptoms.

(3) This percentage represents the percentage of providers that carried out this task on at least one of their patients

(4) These figures are for those children with an under five card

*P value is < 0.001

Table A-3												
Comparison of Observation, Exit Interview, Record Review and Provider Interview for Cough												
COUGH	Obs-Exit Interview N=209				Obs-Record Review N=150				Obs-Provider Interview (1) N=30			
Tasks to be carried out	%obs	% exit	Agree	Kappa	%obs	% rec	Agree	Kappa	% obs (3)	% int	Agree	Kappa
Ask about duration of cough	76%	71%	79%	0.464*	76%	1%	26%	0.009	93%	77%	77%	0.132
Ask about difficult breathing	33%	28%	72%	0.347*	34%	1%	67%	0.026	63%	67%	67%	0.318+
Count respiration rate	10%	16%	82%	0.231*	11%	0%	89%	0.000	17%	53%	43%	-0.085
Observe for chest indrawing	47%	53%	77%	0.5501	48%	0%	52%	0.000	70%	67%	63%	0.154
Check for neck stiffness	3%	4%	94%	-0.031	1%	0%	99%	0.000	7%	13%	80%	-0.098
Classify child by severity of illness	37%	20%	59%	0.027	39%	12%	64%	0.129	60%	40%	60%	0.231
Tell diagnosis of cough	78%	16%	48%	0.114*	data not available for comparison				data not available for comparison			
Give appropriate "drug" for cough	57%	53%	88%	0.7551	63%	67%	90%	0.7801	data not available for comparison			
Give correct cough RX dosages	55%	49%	86%	0.7181	61%	65%	89%	0.754*	data not available for comparison			
Tell to feed during illness	31%	46%	72%	0.416*	35%	1%	67%	0.027	57%	57%	53%	0.049
Tell to feed after illness	24%	36%	66%	0.222*	25%	0%	76%	0.000				
Tell to increase fluid intake	34%	40%	70%	0.369*	39%	1%	59%	-0.069	50%	60%	57%	0.133
Advise to soothe and relieve cough	8%	9%	88%	0.238	9%	1%	91%	0.112	23%	30%	67%	0.153
Tell to return if breathing is difficult	16%	26%	83%	0.496*	21%	0%	79%	0.000	30%	77%	47%	0.121
Tell to return if child unable to drink	7%	17%	81%	0.093	8%	0%	92%	0.000	10%	50%	53%	0.067
Tell to return if child is more ill	21%	35%	77%	0.436*	28%	0%	72%	0.000	37%	63%	60%	0.253+
Tell to return if child has convulsion	3%	15%	84%	0.008	3%	0%	97%	0.0001	data not available for comparison			
Tell to return in 2 days to reassess (2)	150%	24%	68%	0.368	53%	0%	47%	0.0001	data not available for comparison			

(1) In provider interviews, a probed YES was counted as NO

(2) For cases of suspected pneumonia

(3) This percentage represents only those cases observed that were also found in the record review

(4) This percentage represents the percentage of providers that carried out this task on at least one of their patients

*P value is < 0.001

Table A-4												
Comparison of Observation, Exit Interview, Record Review and Provider Interview for Diarrhoea												
DIARRHOEA Tasks to be carried out	Obs-Exit Interview N=123				Obs-Record Review N=86				Obs-Provider Interview (1) N=20			
	%obs	% exit	Agree	Kappa	%obs	% rec	Agree	Kappa	% obs (3)	%	Agree	Kappa
Ask about duration of diarrhoea	78%	83%	87%	0.567*	84%	1%	17%	0.005	90%	77%	79%	0.298+
Ask about blood in stools	49%	39%	74%	0.477*	55%	7%	52%	0.127	69%	60%	62%	0.188
Ask whether urinating	2%	15%	85%	0.056	2%	0%	98%	0.000	7%	14%	79%	-0.105
Pinch skin	28%	24%	71%	0.241*	29%	0%	71%	0.000	52%	90%	62%	0.220+
Check for sunken eyes	58%	46%	61%	0.228*	62%	0%	38%	0.0001	69%	73%	55%	-0.080
Check fontanel (N = 61)	14%	15%	75%	0.390*	21%	0%	76%	0.000	30%	60%	50%	0.083
Classify degree of dehydration	34%	8%	64%	0.026	34%	13%	72%	0.263	46%	26%	46%	-0.125
Treat dehydration appropriately	78%	22%	29%	-0.080	80%	55%	59%	0.183	62%	77%	69%	0.337+
Tell diagnosis of diarrhoea	74%	32%	49%	0.133	no data available for comparison							
Give ORS	83%	81%	91%	0.617*	89%	75%	75%	0.178	see treat dehydration appropriately			
Diagnose normal diarrhoea	86%	86%	98%	0.922*	86%	87%	99%	0.945*	no data available for comparison			
Give appropriate "drug" for diarrhoea	79%	78%	97%	0.906*	82%	87%	95%	0.958*	see treat dehydration appropriately			
Give correct diarrhoea RX dosages	79%	76%	94%	0.843*	82%	87%	95%	0.958*	see treat dehydration appropriately			
Tell to feed during illness	44%	53%	83%	0.6611	48%	0%	52%	0.000	50%	55%	71%	0.429+
Tell to feed after illness	37%	40%	72%	0.415*	41%	0%	59%	0.000	41%	46%	67%	0.319+
Tell to increase fluid intake	48%	59%	78%	0.564*	52%	2%	50%	0.040	64%	69%	82%	0.602+
Explain how to prepare ORS	62%	78%	69%	0.286*	65%	0%	35%	0.000	64%	52%	68%	0.343+
Explain how to administer ORS	64%	79%	71%	0.298*	67%	0%	33%	0.000	71%	45%	68%	0.376+
Tell to return in 3 days if not better	15%	26%	78%	0.343*	17%	0%	83%	0.000	29%	52%	54%	0.099
Tell to return if signs of dehydration	18%	29%	78%	0.393*	21%	0%	79%	0.000	25%	62%	50%	0.096
Tell how to prevent getting diarrhoea	12%	7%	87%	0.226*	13%	0%	87%	0.000	18%	82%	18%	0.0001

(1) In provider interviews, a probed YES was counted as NO (only spontaneous answers)

(2) This percentage represents the percentage of providers that carried out this task on at least one of their patients

*P value is < 0.001

+P value is < 0.05

Table A-5												
Comparison of Observation, Exit Interview, Record Review and Provider Interview for Fever/Malaria												
FEVER/MALARIA Tasks to be carried out	Obs-Exit Interview N=200				Obs-Record Review N=160				Obs-Provider Interview N=29			
	%obs	% exit	Agree	Kappa	%obs	% rec	Agree	Kappa	% obs (3)	%	Agree	Kappa
Ask about presence of other symptoms	70%	85%	68%	0.111+	69%	10%	39%	0.066+	55%	38%	55%	0.125
Ask about vomiting	44%	33%	72%	0.400*	41%	0%	60%	0.000	93%	66%	66%	0.058
Ask about treatment already taken	55%	41%	74%	0.491*	57%	1%	42%	-0.013	76%	55%	59%	0.126
Take temperature	43%	45%	90%	0.796*	45%	0%	55%	0.000	52%	79%	52%	0.015
Check for splenomegaly	28%	30%	76%	0.410*	25%	0%	75%	0.000	41%	41%	52%	0.004
Take blood for a slide	1%	2%	99%	0.496*	0%	0%	100%	0.000	4%	18%	79%	-0.063
Check for anemia	63%	55%	70%	0.384*	62%	0%	38%	0.000	69%	52%	62%	0.231
Give S-P	80%	77%	95%	0.867*	87%	88%	97%	0.860*	no data available for comparison			
Ask to wait for 30 minutes	23%	8%	97%	0.374*	23%	0%	73%	0.000	47%	27%	67%	0.312+
Tell diagnosis of malaria	81%	38%	53%	0.180*	no data available for comparison							
Give appropriate "drug" for malaria	94%	93%	97%	0.683*	94%	93%	98%	0.813*	97%	93%	90%	-0.046
Give correct malaria RX dosages	94%	91%	95%	0.614*	93%	93%	98%	0.831*	97%	83%	80%	-0.060
Tell to feed during illness	36%	35%	74%	0.432*	41%	0%	58%	0.000	53%	50%	63%	0.267
Tell to feed after illness	30%	28%	75%	0.391*	34%	0%	58%	0.000	37%	30%	67%	0.254
Tell to increase fluid intake	36%	38%	71%	0.380*	40%	4%	63%	0.113*	53%	63%	57%	0.117
Advise to give child sponge bath	25%	28%	75%	0.371*	29%	3%	72%	0.081+	50%	70%	41%	-0.114
Tell to return in 3 days if not better	26%	29%	73%	0.309*	29%	0%	70%	0.000	43%	63%	53%	0.098
Tell how to prevent malaria	1%	2%	99%	0.496*	5%	0%	99%	0.000	3%	90%	13%	0.008

(1) In provider interviews, a probed YES was counted as NO (only spontaneous answers)

(2) This percentage represents only those cases observed that were also found during the record review.

(3) This percentage represents the percentage of providers that carried out this task on at least one of their patients

*P value is < 0.001

+P value is < 0.05

Table A-6				
Comparison of Observation and Exit Interview Data for Integrated Management of Children with Fever and Diarrhoea				
FEVER and DIARRHOEA	Obs-Exit Interview (N = 54)			
Tasks to be carried out	% obs	% exit	% Agree	Kappa
Ask \geq 1 history question for each problem	65%	59%	72%	0.412*
Ask \geq 2 history questions for each problem	24%	17%	85%	0.547*
Asked all history questions for each problem	0%	4%	96%	—
Did \geq 1 physical exam for each problem	43%	48%	72%	0.441
Did \geq 2 physical exams for each problem	19%	11%	81%	0.274+
Did all physical exams for each problem	0%	0%	100%	—
Treated correctly for both problems	83%	67%	80%	0.476*
Gave \geq 1 advice on feeding	28%	33%	80%	0.522*
Gave all on feeding	24%	26%	80%	0.457*
Gave advice on home care for each problem	0%	0%	100%	—
Gave any advice on when to return for each	9%	19%	83%	0.315+

* P value is < 0.001

+ P value is < 0.05

Table A-7				
Comparison of Observation and Exit Interview Data for Integrated Management of Children with Cough and Fever				
COUGH and FEVER	Obs-Exit Interview (N = 54)			
Tasks to be carried out	% obs	% exit	% Agree	Kappa
Ask ≥ 1 history question for each problem	73%	48%	64%	0.293*
Ask ≥ 2 history questions for each problem	33%	21%	75%	0.374*
Asked all history questions for each problem	1%	1%	97%	-0.013
Did ≥ 1 physical exam for each problem	43%	39%	72%	0.421 *
Did ≥ 2 physical exams for each problem	4%	7%	92%	0.211+
Did all physical exams for each problem	0%	0%	100%	—
Treated correctly for both problems	49%	45%	77%	0.546*
Gave ≥ 1 advice on feeding	23%	17%	81%	0.419*
Gave all on feeding	4%	0%	96%	—
Gave advice on home care for each problem	3%	0%	97%	—
Gave any advice on when to return for each	8%	11%	84%	0.056

* P value is < 0.001

+ P value is < 0.05

Table A-8			
Consistency of Provider Performance Related to Cough for 30 Providers Observed Managing 3 or more Cough Patients			
COUGH	Providers Observed = 30		
Tasks to be carried out	% ever doing	% consistent (always or never)	% always doing if ever did
Ask about duration of cough	93%	50%	46%
Ask about difficult breathing	53%	53%	13%
Count respiration rate	17%	83%	0%
Observe for chest indrawing	76%	33%	13%
Check for neck stiffness	7%	93%	0%
Classify child by severity of illness	63%	50%	21%
Tell to feed during illness	53%	53%	13%
Tell to feed after illness	37%	67%	9%
Tell to increase fluid intake	27%	50%	6%
Advise to soothe and relieve cough	27%	73%	0%
Tell to return if breathing is difficult	37%	63%	0%
Tell to return if child unable to drink	10%	90%	0%
Tell to return if child is more ill	40%	63%	8%
Tell to return if child has convulsion	7%	93%	0%
Tell diagnosis of cough	87%	50%	42%
Give appropriate "drug" for cough	93%	23%	18%
Give correct cough RX dosages	93%	23%	18%

Table A-9			
Consistency of Provider Performance Related to Diarrhoea for 20 Providers Observed Managing 3 or more Diarrhoea Patients			
DIARRHOEA	Providers Observed = 20		
Tasks to be carried out	% ever doing	% consistent (always or never)	% always doing if ever did
Ask about duration of diarrhoea	95%	65%	63%
Ask about blood in stools	80%	30%	13%
Ask whether urinating	15%	85%	0%
Pinch skin	60%	50%	17%
Check for sunken eyes	80%	40%	25%
Classify degree of dehydration	55%	65%	36%
Treat dehydration appropriately	90%	50%	44%
Tell to feed during illness	55%	70%	45%
Tell to feed after illness	45%	70%	33%
Tell to increase fluid intake	60%	65%	42%
Explain how to prepare ORS	80%	55%	44%
Explain how to administer ORS	75%	60%	47%
Tell to return in 3 days if not better	30%	85%	50%
Tell to return if signs of dehydration	35%	70%	14%
Tell how to prevent getting diarrhoea	25%	80%	20%
Tell diagnosis of diarrhoea	90%	80%	78%
Give appropriate "drug" for diarrhoea	95%	50%	47%
Give correct diarrhoea RX dosages	95%	50%	47%

Table A-1 0			
Consistency of Provider Performance Related to Fever for 25 Providers Observed Managing 3 or more Fever Patients			
FEVER/MALARIA	Providers Observed = 25		
Tasks to be carried out	% ever doing	% consistent (always or never)	% always doing if ever did
Ask about presence other symptoms	52%	68%	38%
Ask about vomiting	96%	32%	29%
Ask about treatment already taken	88%	32%	23%
Take temperature	64%	54%	25%
Check for splenomegaly	48%	60%	70%
Take blood for a slide	0%	100%	0%
Check for anemia	72%	52%	33%
Ask to wait for 30 minutes	52%	52%	8%
Tell to feed during illness	56%	50%	29%
Tell to feed after illness	40%	72%	30%
Tell to increase fluid intake	52%	60%	23%
Advise to give child sponge bath	56%	44%	0%
Tell to return in 3 days if not better	52%	60%	23%
Tell how to prevent malaria	4%	96%	0%
Tell diagnosis of malaria	84%	48%	38%
Give appropriate "drug" for malaria	100%	68%	68%
Give correct malaria RX dosages	100%	64%	64%

Appendix B

Data Collection Instruments

- **Observation: Treatment of Children Under Five**
- **Exit Interview with Mothers: Treatment of Children Under Five (English Version)**
- **Record Review: Treatment of Children Under Five**
- **Interview with Health Care Provider**
- **Time Requirements for Observations**

- 1 At the time of the study, there were no patient register at government facilities in the rest of Malawi, only a system for counting the number of patients according to age group and diagnosis. The patient register is now being introduced at a national level.
- 2 Medical assistants receive 2-3 years training post-high school, while clinical officers receive 3-4 years training. Medical assistants generally are responsible for work in the OPD, while clinical officers also have surgical and in-patient responsibilities.
- 3 These are booklets prepared by the Ministry of Health and distributed to all health workers. They were developed with support from the Malawi Essential Drug Programme, and provide standards for history-taking, physical examination, diagnosis, treatment, counseling, and follow-up for common conditions.
- 4 It was not possible to devise a weighted score, since the number of cases seen by each provider of each main symptom were not the same. For example, in the case of diarrhoea, some providers were only observed seeing one case, while others saw as many as 15 cases.

-
- 5 Because data collectors were present at the clinics for several days and the clinics were often quite busy, the investigators decided not to allow observers to interrupt the providers' work by asking questions about what they were doing.
 - 6 No comparison for made with provider interviews because the interview questions focused only on treatment of pneumonia and there were too few observed pneumonia cases.
 - 7 When the proportions are very large, Kappa is low as the chance of not agreeing would be small anyway.